

# An Entropy Stable Discontinuous Galerkin Finite-Element Moment Method for the Boltzmann Equation

M.R.A. Abdelmalik\*, E.H. van Brummelen

*Department of Mechanical Engineering, Eindhoven University of Technology, The Netherlands*

---

## Abstract

This paper presents a numerical approximation technique for the Boltzmann equation based on a moment-system approximation in velocity dependence and a discontinuous Galerkin finite-element approximation in position dependence. The closure relation for the moment systems derives from minimization of a suitable  $\varphi$ -divergence. This divergence-based closure yields a hierarchy of tractable symmetric hyperbolic moment systems that retain the fundamental structural properties of the Boltzmann equation. The resulting combined discontinuous Galerkin moment method corresponds to a Galerkin approximation of the Boltzmann equation in renormalized form. We present a new class of numerical flux functions, based on the underlying renormalized Boltzmann equation, that ensure entropy dissipation of the approximation scheme. Numerical results are presented for a one-dimensional test case.

**Keywords:** Boltzmann equation, kinetic theory, moment systems, hyperbolic systems, entropy,  $\varphi$ -divergence, discontinuous Galerkin finite-element methods, numerical flux functions

---

## 1. Introduction

The Boltzmann equation provides a description of the molecular dynamics of fluid flows based on their one-particle phase-space distribution. However, the Boltzmann equation also encapsulates all conventional macroscopic flow models in the sense that its limit solutions correspond to solutions of the compressible Euler and Navier–Stokes equations [1, 2], the incompressible Euler and Navier–Stokes equations [3, 4], the incompressible Stokes equations [5] and the incompressible Navier–Stokes–Fourier system [6]; see [7] for an overview. The Boltzmann equation is uniquely suited to describe flows in the transitional molecular/continuum regime and the corresponding rarefaction effects, by virtue of its inherent characterization of deviations of the velocity distribution from local equilibrium. Applications in which rarefaction effects play a significant role are multitudinous, including gas flow problems involving large mean free paths in high-altitude flows and hypobaric applications such as chemical vapor deposition; see [8, 9] and references therein for further examples. Moreover, the perpetual trend toward miniaturization in science and technology renders accurate descriptions of fluid flows in the transitional molecular/continuum regime of fundamental technological relevance, for instance, in nanoscale applications, micro-channel flows or flow in porous media [10]. The Boltzmann equation also provides a prototype for kinetic models in many other applications that require a description of the collective behavior of large ensembles of small particles, for instance, in semi-conductors [11], in plasmas and fusion and fission devices [12] and in dispersed-particle flows such as in fluidized-bed reactors [13, 14, 15].

Numerical approximation of the Boltzmann equation poses a formidable challenge, on account of its high dimensional setting: for a problem in  $D$  spatial dimensions, the one-particle phase-space is  $2D$  dimensional. The corresponding computational complexity of conventional discretization methods for (integro-)differential equations, such as finite-element methods with uniform meshes, is prohibitive. Numerical approximations of the Boltzmann equation have been predominantly based on particle methods, such as the Direct Simulation Monte Carlo (DSMC) method [16, 17]. Convergence proofs for these methods [18] however convey that their computational complexity depends sensitively on the Knudsen number, and the computational cost becomes

---

\*Corresponding author

Email address: m.abdel.malik@tue.nl (E.H. van Brummelen)

prohibitive in the fluid-dynamical limit. Moreover, from an approximation perspective, DSMC can be inefficient, because it is inherent to the underlying Monte-Carlo process that the approximation error decays only as  $n^{-\frac{1}{2}}$  as the number of simulation molecules,  $n$ , increases; see, for instance, [19, Thm. 5.14]. Efficient computational modeling of fluid flows in the transitional molecular/continuum regime therefore remains an outstanding challenge.

An alternative approximation technique for the Boltzmann equation is the method of moments [20, 21, 9]. The method of moments represents a general statistical approximation technique which identifies parameters of an approximate distribution based on its moments [22]. Application of the method of moments to the Boltzmann equation engenders an evolution equation for the moments (weighted averages) of the phase-space distribution. An approximation based on moments is generally consistent with a restricted interest in functionals of the distribution corresponding to macroscopic properties of the fluid. The method of moments is closely related to extended thermodynamics; see, for instance, [23, 24].

In this paper we present a position-velocity Galerkin approximation method for the Boltzmann equation, based on a moment-system approximation in velocity dependence and a discontinuous Galerkin approximation in position dependence. To devise the moment-closure relation, we consider a generalization of the setting of the moment-closure problem from Kullback–Leibler divergence [25] to the class of  $\varphi$ -divergences [26, 27]. The derived moment-closure relation engenders non-negative distributions and the corresponding moment systems are symmetric hyperbolic and tractable, in the sense that the formulation only requires the evaluation of higher-order moments of Gaussian distributions. The moment systems conserve mass, momentum and energy, and moreover dissipate an appropriate  $\varphi$ -divergence, analogous to the dissipation of relative entropy of the underlying Boltzmann equation, provided that the collision operator dissipates the corresponding  $\varphi$ -divergence relative to a suitable reference distribution. The moment systems correspond to Galerkin approximations of the Boltzmann equation in renormalized form. For the discretization in position dependence, we consider a discontinuous Galerkin finite-element method (DGFEM). We present a new numerical-flux function for the DGFEM discretization, derived from the underlying renormalized Boltzmann equation. We will show that the resulting DGFEM moment method is entropy stable.

The remainder of this paper is organized as follows. Section 2 surveys standard structural properties of the Boltzmann equation that have to be retained in the moment-system approximation. Section 3 introduces the moment-system approximation of the Boltzmann equation in velocity dependence and the corresponding moment-closure relation. In Section 4, we derive the DGFEM formulation in position dependence, and we show that the resulting DGFEM moment method is entropy stable. Section 5 illustrates the properties of the proposed DGFEM moment method for a shock tube Riemann problem [28, 29]. Finally, Section 6 presents a concluding discussion.

## 2. Properties of the Boltzmann equation

Consider a monatomic gas contained in a fixed spatial domain  $\Omega \subset \mathbb{R}^D$ . Kinetic theory describes the state of such a gas by a non-negative (phase-space) density  $f = f(t, \mathbf{x}, \mathbf{v})$  over the single-particle phase space  $\Omega \times \mathbb{R}^D$ . The evolution of  $f$  is considered to be governed by the Boltzmann equation,

$$\partial_t f + v_i \partial_{x_i} f = \mathcal{C}(f) \quad (1)$$

where the summation convention applies to repeated indices and the collision operator  $f \mapsto \mathcal{C}(f)$  acts only on the  $\mathbf{v} = (v_1, \dots, v_D)$  dependence of  $f$  locally at each  $(t, \mathbf{x})$ . The collision operator is assumed to possess certain conservation, symmetry and dissipation properties, viz., conservation of mass, momentum and energy, invariance under Galilean transformations and dissipation of appropriate entropy functionals. Moreover, it is assumed that the collision operator exhibits certain positivity properties. These fundamental properties of the collision operator are treated in further detail below. Our treatment of the conservation and symmetry properties is standard (see, for instance, [21]) and is presented merely for coherence and completeness. For the entropy-dissipation property, we consider a generalization of the usual (relative) entropy to  $\varphi$ -divergences [26, 27] for collision operators  $f \mapsto \mathcal{C}(f)$  that are non-negative outside the support of  $f$ , to enable an exploration of the moment-closure problem in an extended setting.

To elaborate the conservation properties of the collision operator, let  $\langle \cdot \rangle$  denote integration in the velocity dependence of any scalar, vector or matrix valued measurable function over  $D$ -dimensional Lebesgue measure. A function  $\alpha : \mathbb{R}^D \rightarrow \mathbb{R}$  is called a *collision invariant* of  $\mathcal{C}$  if

$$\langle \alpha \mathcal{C}(f) \rangle = 0 \quad \forall f \in \mathcal{D}(\mathcal{C}), \quad (2)$$

where  $\mathcal{D}(\mathcal{C}) \subset L^1(\mathbb{R}^D, \mathbb{R}_{\geq 0})$  denotes the domain of  $\mathcal{C}$ , which we consider to be a subset of the almost everywhere nonnegative Lebesgue integrable functions on  $\mathbb{R}^D$ . Equation (1) associates a scalar conservation law with each collision invariant:

$$\partial_t \langle \alpha f \rangle + \partial_{x_i} \langle v_i \alpha f \rangle = 0 \quad (3)$$

We insist that  $\{1, v_1, \dots, v_D, |\mathbf{v}|^2\}$  are collision invariants of  $\mathcal{C}$  and that the span of this set contains all collision invariants, i.e.

$$\langle \alpha \mathcal{C}(f) \rangle = 0 \quad \forall f \in \mathcal{D}(\mathcal{C}) \quad \Leftrightarrow \quad \alpha \in \text{span}\{1, v_1, \dots, v_D, |\mathbf{v}|^2\} =: \mathcal{I}. \quad (4)$$

The moments  $\langle f \rangle$ ,  $\langle v_i f \rangle$  and  $\langle |\mathbf{v}|^2 f \rangle$ , correspond to mass-density, the (components of) momentum-density and energy-density, respectively. Accordingly, the conservation law (3) implies that (1) conserves mass, momentum and energy.

The assumed symmetry properties of the collision operator pertain to commutation with translational and rotational transformations. In particular, for all vectors  $\mathbf{u} \in \mathbb{R}^D$  and all orthogonal tensors  $\mathcal{O} : \mathbb{R}^D \rightarrow \mathbb{R}^D$ , we define the translation transformation  $\mathcal{T}_{\mathbf{u}} : \mathcal{D}(\mathcal{C}) \rightarrow \mathcal{D}(\mathcal{C})$  and the rotation transformation  $\mathcal{T}_{\mathcal{O}} : \mathcal{D}(\mathcal{C}) \rightarrow \mathcal{D}(\mathcal{C})$  by:

$$(\mathcal{T}_{\mathbf{u}} f)(\mathbf{v}) = f(\mathbf{u} - \mathbf{v}) \quad \forall f \in \mathcal{D}(\mathcal{C}) \quad (5)$$

$$(\mathcal{T}_{\mathcal{O}} f)(\mathbf{v}) = f(\mathcal{O}^* \mathbf{v}) \quad \forall f \in \mathcal{D}(\mathcal{C}) \quad (6)$$

with  $\mathcal{O}^*$  the Euclidean adjoint of  $\mathcal{O}$ . Note that the above transformations act on the  $\mathbf{v}$ -dependence only. It is assumed that  $\mathcal{C}$  possesses the following symmetries:

$$\mathcal{C}(\mathcal{T}_{\mathbf{u}} f) = \mathcal{T}_{\mathbf{u}} \mathcal{C}(f), \quad \mathcal{C}(\mathcal{T}_{\mathcal{O}} f) = \mathcal{T}_{\mathcal{O}} \mathcal{C}(f) \quad (7)$$

The symmetries (7) imply that (1) complies with Galilean invariance, i.e. if  $f(t, \mathbf{x}, \mathbf{v})$  satisfies the Boltzmann equation (1), then for arbitrary  $\mathbf{u} \in \mathbb{R}^D$  and arbitrary orthogonal  $\mathcal{O} : \mathbb{R}^D \rightarrow \mathbb{R}^D$ , so do  $f(t, \mathbf{x} - \mathbf{u}t, \mathbf{v} - \mathbf{u})$  and  $f(t, \mathcal{O}^* \mathbf{x}, \mathcal{O}^* \mathbf{v})$ .

The entropy dissipation property of  $\mathcal{C}$  is considered in the extended setting of [21, Sec. 7], from which we derive the following definition: a convex function  $\eta : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$  is called an *entropy for  $\mathcal{C}$*  if

$$\langle \eta'(f) \mathcal{C}(f) \rangle \leq 0, \quad \forall f \in \mathcal{D}(\mathcal{C}) \quad (8)$$

with  $\eta'(f)$  the derivative of  $\eta(f)$ , and if for every  $f \in \mathcal{D}(\mathcal{C})$  the following equivalences hold:

$$\mathcal{C}(f) = 0 \quad \Leftrightarrow \quad \langle \eta'(f) \mathcal{C}(f) \rangle = 0 \quad \Leftrightarrow \quad \eta'(f) \in \mathcal{I} \quad (9)$$

Relation (8) implies that  $\mathcal{C}$  dissipates the local entropy density  $\langle \eta(\cdot) \rangle$ , which leads to an abstraction of Boltzmann's H-theorem for (1), asserting that solutions of the Boltzmann equation (1) satisfy the local entropy-dissipation law:

$$\partial_t \langle \eta(f) \rangle + \partial_{x_i} \langle v_i \eta(f) \rangle = \langle \mathcal{C}(f) \eta'(f) \rangle \leq 0. \quad (10)$$

The functions  $\langle \eta(f) \rangle$ ,  $\langle v_i \eta(f) \rangle$  and  $\langle \eta'(f) \mathcal{C}(f) \rangle$  are referred to as entropy density, entropy flux and entropy-dissipation rate, respectively. The first equivalence in (9) characterizes local equilibria of  $\mathcal{C}$  by vanishing entropy dissipation, while the second equivalence indicates the form of such local equilibria. For spatially homogeneous initial data,  $f_0$ , equations (8) and (9) suggest that equilibrium solutions,  $f_{\text{eq}}$ , of (1) are determined by:

$$f_{\text{eq}} = \arg \min \{ \langle \eta(f) \rangle : f \in \mathcal{D}(\mathcal{C}), \langle \alpha f \rangle = \langle \alpha f_0 \rangle \quad \forall \alpha \in \mathcal{I} \}, \quad (11)$$

Equation (11) identifies equilibria as minimizers<sup>1</sup> of the entropy, subject to the constraint that the invariant moments are identical to the invariant moments of the initial distribution.

The standard definition of entropy corresponds to a density  $f \mapsto \langle f \log f + f \alpha \rangle$  where  $\alpha \in \mathcal{I}$  is any collision invariant. The corresponding local equilibria of  $\mathcal{C}(f)$  defined by (9) are characterized by Maxwellians  $\mathcal{M}$ , i.e. distributions of the form

$$\mathcal{M}(\mathbf{v}) := \mathcal{M}_{(\varrho, \mathbf{u}, T)}(\mathbf{v}) := \frac{\varrho}{(2\pi RT)^{\frac{D}{2}}} \exp\left(-\frac{|\mathbf{v} - \mathbf{u}|^2}{2RT}\right) \quad (12)$$

---

<sup>1</sup>We adopt the sign convention of diminishing entropy.

for some  $(\varrho, \mathbf{u}, T) \in \mathbb{R}_{>0} \times \mathbb{R}^D \times \mathbb{R}_{>0}$  and a certain gas constant  $R \in \mathbb{R}_{>0}$ . It is noteworthy that  $\log \mathcal{M} \in \mathcal{I}$  and, therefore, the Kullback-Leibler divergence  $\langle f \log(f/\mathcal{M}) \rangle$  of  $f$  with respect to  $\mathcal{M}$  is equivalent to  $\langle f \log f \rangle$  in the sense of dissipation characteristics.

In this work, we depart from the standard relative entropy for (1) corresponding to the Kullback-Leibler divergence. Instead, we consider entropies based on particular  $\varphi$ -divergences [30, 26, 27]. These  $\varphi$ -divergences generally preclude the usual physical and information-theoretical interpretations of relative entropy, but still provide a meaningful entropy density in accordance with (8) and (9). The considered  $\varphi$ -divergences yield a setting in which entropy-minimization based moment-closure approximations to (1) are not impaired by non-realizability, exhibit bounded fluxes in the vicinity of equilibrium, and are numerically tractable.

We will admit distributions that vanish on sets with nonzero measure. To accommodate such distributions, we introduce an auxiliary non-negativity condition on the collision operator, in addition to (8) and (9). The non-negativity condition insists that  $\mathcal{C}(f)$  cannot be negative on zero sets of  $f$ :

$$\mathcal{C}(f)|_{\text{supp}^c(f)} \geq 0 \quad (13)$$

where  $\text{supp}^c(f)$  denotes zero set of  $f$ , i.e. the complement in  $\mathbb{R}^D$  of the closed support of  $f$ . Condition (13) encodes that the collision operator cannot create locally negative distributions. It can be verified that (13) holds for a wide range of collision operators, including the BGK operator [31], the multi-scale generalization of the BGK operator introduced in [21], and all collision operators that are characterized by a (non-negative) collision kernel.

**Remark.** *The adoption of  $\varphi$ -divergence-based entropies insists that such entropies satisfy (8) and (9) for a meaningful class of collision operators in compliance with (13). In [27] it is shown that the class of admissible collision operators includes the BGK operator [31] and the multi-scale generalization of the BGK operator in [21].*

### 3. Moment-system approximation of the Boltzmann equation

Our semi-discretization of the Boltzmann equation with respect to the velocity dependence is based on velocity moments of the one-particle marginal. An inherent aspect of considering a finite number of moment equations derived from (1) is that low-order moments are generally coupled to higher-order ones. Consequently, a closed set of equations for the moments cannot be readily formulated, and a closure relation is required. Closed moment systems can generally be conceived of as Galerkin-subspace approximations of (1) in renormalized form.

To derive the moment equations from (1), let  $\mathcal{M}$  denote a finite-dimensional subspace of  $D$ -variate polynomials and let  $\{m_i(\mathbf{v})\}_{i=1}^M$  with  $M = \dim \mathcal{M}$  represent a corresponding basis. Denoting the column  $M$ -vector of these basis elements by  $\mathbf{m}$ , it holds that the moments  $\langle \mathbf{m}f \rangle$  of the one-particle marginal satisfy:

$$\partial_t \langle \mathbf{m}f \rangle + \partial_{x_i} \langle v_i \mathbf{m}f \rangle = \langle \mathbf{m} \mathcal{C}(f) \rangle \quad (14)$$

provided that  $f \in \mathcal{F} := \{f \in \mathcal{D}(\mathcal{C}) : f \geq 0, m f \in L^1(\mathbb{R}^D), \mathbf{v} m f \in L^1(\mathbb{R}^D, \mathbb{R}^D), m \mathcal{C}(f) \in L^1(\mathbb{R}^D) \forall m \in \mathcal{M}\}$  almost everywhere in the considered time interval  $(0, T)$  and the spatial domain  $\Omega$ . The provision  $f \in \mathcal{F}$  has been confirmed in specific settings of (1) but not for the general case; see [21, Sec. 4]. The moment system (14) constitutes  $M$  relations between  $(2 + D) \mathbb{R}^M$ -valued functions, viz. the density  $\langle \mathbf{m}f \rangle$ , the fluxes  $\langle v_i \mathbf{m}f \rangle$  and the production term  $\langle \mathbf{m} \mathcal{C}(f) \rangle$ , and is therefore not closed. Moment systems are generally closed by constructing an approximation to the distribution function from the densities and then evaluating the fluxes and production terms for the approximate distribution. Denoting by  $\mathbb{M} \subseteq \mathbb{R}^M$  a suitable class of moments, a function  $\mathcal{F} : \mathbb{M} \rightarrow \mathcal{F}$  must be specified such that  $\mathcal{F}$  realizes the moments in  $\mathbb{M}$ , i.e.  $\langle \mathbf{m} \mathcal{F}(\boldsymbol{\mu}) \rangle = \boldsymbol{\mu}$  for all  $\boldsymbol{\mu} \in \mathbb{M}$ , and if  $f$  satisfies (1) then  $\mathcal{F}(\langle \mathbf{m}f \rangle)$  constitutes a suitable (in a sense to be made more precise below) approximation to  $f$ . Approximating the moments in (14) by  $\boldsymbol{\mu} \approx \langle \mathbf{m}f \rangle$  and replacing  $f$  in (14) by the approximation  $\mathcal{F}(\boldsymbol{\mu})$ , one obtains the following closed system for the approximate moments:

$$\partial_t \boldsymbol{\mu} + \partial_{x_i} \langle v_i \mathbf{m} \mathcal{F}(\boldsymbol{\mu}) \rangle = \langle \mathbf{m} \mathcal{C}(\mathcal{F}(\boldsymbol{\mu})) \rangle. \quad (15)$$

The closed moment system (15) is essentially defined by the polynomial subspace,  $\mathcal{M}$ , and the closure relation,  $\mathcal{F}$ . A subspace/closure-relation pair  $(\mathcal{M}, \mathcal{F})$  is appropriate if the corresponding moment system (15) is well posed and retains the fundamental structural properties of the Boltzmann equation: conservation of mass, momentum and energy, Galilean invariance and dissipation of an entropy functional. Further conditions may be considered,

e.g. that the fluxes and production terms can be efficiently evaluated by means of numerical quadrature. The conservation properties and Galilean invariance of (1) can generally be transferred to (15) by a suitable selection of the subspace  $\mathcal{M}$ , namely that  $\mathcal{M}$  contains the collision invariants  $\mathcal{I}$ , and is closed under the actions of  $\mathcal{T}_u$  and  $\mathcal{T}_O$  (cf. (5) and (6)). Entropy dissipation must be ensured by the closure relation,  $\mathcal{F}$ .

We consider a moment-closure relation deriving from minimization of a  $\varphi$ -divergence relative to a suitable background measure, subject to the moment constraints; see also [27]. An example of the class of  $\varphi$ -divergences that we envisage is provided by:

$$\varphi(s) = s \left( \frac{N^2}{N+1} s^{1/N} - N \right) + \frac{N}{N+1} \quad (16)$$

with  $N \in \mathbb{N}$ . We assume that the considered  $\varphi$ -divergence is strictly convex and represents an entropy for the collision operator. We denote by  $\mathcal{B}$  a strictly positive background distribution. We assume that  $\mathcal{B}$  is independent of  $(t, \mathbf{x})$ . However, this assumption can be moderated. We consider the closure relation  $\mathcal{F} : \mathbb{M} \rightarrow \mathcal{F}$  according to:

$$\mathcal{F}(\boldsymbol{\mu}) = \arg \min_{f \in \mathcal{F}} \{ \langle \mathcal{B} \varphi(f/\mathcal{B}) \rangle : f \in \mathcal{D}(\mathcal{C}), \langle \mathbf{m} f \rangle = \boldsymbol{\mu} \} \quad (17)$$

Assuming that (17) admits a solution (see [27]), this solution can be characterized as a stationary point of the Lagrangian  $(f, \boldsymbol{\lambda}) \mapsto \langle \mathcal{B} \varphi(f/\mathcal{B}) \rangle + \boldsymbol{\lambda} \cdot (\boldsymbol{\mu} - \langle \mathbf{m} f \rangle)$ , where  $\boldsymbol{\lambda}$  represents a Lagrange multiplier. The stationarity condition implies  $\varphi'(f/\mathcal{B}) - \boldsymbol{\lambda} \cdot \mathbf{m} = 0$  and, hence, the moment-closure relation:

$$\mathcal{F}(\boldsymbol{\mu}) = \mathcal{B} \psi(\boldsymbol{\lambda} \cdot \mathbf{m}) \quad (18)$$

with  $\psi$  a suitable right inverse of  $\varphi'$ . In particular, we define

$$\psi(s) = \begin{cases} (\varphi')^{-1}(s) & \text{if } s > \varphi'(0) \\ 0 & \text{if } s \leq \varphi'(0) \end{cases} \quad (19)$$

By virtue of the strict convexity of  $\varphi$ , the derivative  $\varphi'$  is an increasing function and the inverse in (19) is uniquely defined. It is important to note that  $\varphi'$  corresponds to a left inverse of  $\psi$  on the support of  $\mathcal{F}(\boldsymbol{\mu})$ , i.e. on the closure of  $\{\mathbf{v} \in \mathbb{R}^D : (\varphi')^{-1}(\boldsymbol{\lambda} \cdot \mathbf{m}(\mathbf{v})) > 0\}$ :

$$\varphi'(\psi(\boldsymbol{\lambda} \cdot \mathbf{m}(\mathbf{v}))) = \boldsymbol{\lambda} \cdot \mathbf{m}(\mathbf{v}) \quad \text{for all } \mathbf{v} \in \text{cl}(\{\mathbf{v} \in \mathbb{R}^D : (\varphi')^{-1}(\boldsymbol{\lambda} \cdot \mathbf{m}(\mathbf{v})) > 0\}) \quad (20)$$

For all  $\mathbf{v} \in \text{supp}^c(\mathcal{F}(\boldsymbol{\mu}))$ , it holds that  $\psi(\boldsymbol{\lambda} \cdot \mathbf{m}(\mathbf{v})) = 0$  but generally  $\varphi'(0) \neq \boldsymbol{\lambda} \cdot \mathbf{m}(\mathbf{v})$ . To establish that under suitable boundary conditions, the moment system (15) with closure relation (18) dissipates the relative entropy with density  $f \mapsto \langle \mathcal{B} \varphi(f/\mathcal{B}) \rangle$ , we will show that if  $\mathcal{F} := \mathcal{F}(\boldsymbol{\mu})$  satisfies (15) then

$$\frac{d}{dt} \int_{\Omega} \langle \mathcal{B} \varphi(\mathcal{F}/\mathcal{B}) \rangle + \int_{\partial\Omega} \langle v_n \mathcal{B} \varphi(\mathcal{F}/\mathcal{B}) \rangle \leq 0 \quad (21)$$

where  $v_n = \mathbf{v} \cdot \mathbf{n}$  with  $\mathbf{n}$  the exterior unit normal vector on  $\partial\Omega$ . To this end, we first note that by the chain rule, we have:

$$\frac{d}{dt} \int_{\Omega} \langle \mathcal{B} \varphi(\mathcal{F}/\mathcal{B}) \rangle + \int_{\partial\Omega} \langle v_n \mathcal{B} \varphi(\mathcal{F}/\mathcal{B}) \rangle = \int_{\Omega} (\partial_t \langle \mathcal{B} \varphi(\mathcal{F}/\mathcal{B}) \rangle + \partial_{x_i} \langle v_i \mathcal{B} \varphi(\mathcal{F}/\mathcal{B}) \rangle) = \int_{\Omega} (T_1 + T_2) \quad (22)$$

with

$$\begin{aligned} T_1 &= \langle \varphi'(\mathcal{F}/\mathcal{B}) (\partial_t \mathcal{F} + v_i \partial_{x_i} \mathcal{F}) \rangle \\ T_2 &= \langle (\varphi(\mathcal{F}/\mathcal{B}) - (\mathcal{F}/\mathcal{B}) \varphi'(\mathcal{F}/\mathcal{B})) (\partial_t \mathcal{B} + v_i \partial_{x_i} \mathcal{B}) \rangle \end{aligned}$$

On account of  $\varphi'(\mathcal{F}/\mathcal{B}) = \boldsymbol{\lambda} \cdot \mathbf{m}$  on  $\text{supp}(\mathcal{F})$  and (15), we obtain for the first term:

$$\begin{aligned} T_1 &= \langle (\boldsymbol{\lambda} \cdot \mathbf{m}) (\partial_t \mathcal{F} + v_i \partial_{x_i} \mathcal{F}) \rangle + \int_{\text{supp}^c(\mathcal{F})} (\varphi'(0) - \boldsymbol{\lambda} \cdot \mathbf{m}) (\partial_t \mathcal{F} + v_i \partial_{x_i} \mathcal{F}) \\ &= \langle (\boldsymbol{\lambda} \cdot \mathbf{m}) \mathcal{C}(\mathcal{F}) \rangle + \int_{\text{supp}^c(\mathcal{F})} (\varphi'(0) - \boldsymbol{\lambda} \cdot \mathbf{m}) (\partial_t \mathcal{F} + v_i \partial_{x_i} \mathcal{F}) \\ &= \langle \varphi'(\mathcal{F}/\mathcal{B}) \mathcal{C}(\mathcal{F}) \rangle + \int_{\text{supp}^c(\mathcal{F})} (\varphi'(0) - \boldsymbol{\lambda} \cdot \mathbf{m}) (\partial_t \mathcal{F} + v_i \partial_{x_i} \mathcal{F} - \mathcal{C}(\mathcal{F})) \end{aligned} \quad (23)$$

The first term in the ultimate expression in (23) is non-positive under the standing assumption that  $f \mapsto \langle \mathcal{B}\varphi(f/\mathcal{B}) \rangle$  represents an entropy for the collision operator in accordance with (8). Moreover, on the zero set  $\text{supp}^c(\mathcal{F})$  it holds that  $\varphi'(0) \geq \boldsymbol{\lambda} \cdot \mathbf{m}(\mathbf{v})$ ,  $\partial_t \mathcal{F} + \partial_{x_i} \mathcal{F} = 0$ , and  $\mathcal{C}(\mathcal{F}) \geq 0$ . The latter inequality follows directly from the non-negativity condition (13). Hence,  $T_1$  is non-positive. The second term,  $T_2$ , in the right member of (22) vanishes because the background distribution  $\mathcal{B}$  is independent of  $(t, \mathbf{x})$ .

It is noteworthy that the entropy-dissipation inequality (21) can be extended to non-uniform background distributions. The entropy-dissipation inequality holds, for instance, if the background distribution is a locally Maxwellian flow [20, Appendix 2] or a solution to the Vlasov equation, in which case  $\partial_t \mathcal{B} + v_i \partial_{x_i} \mathcal{B} = 0$ . Let us also allude to the fact that the analysis in (22)-(23) relies on continuity of  $\mathcal{B}\varphi(\mathcal{F}/\mathcal{B})$  in  $(\mathbf{x}, \mathbf{v})$ , for otherwise the chain rule in (22) is invalid.

To demonstrate that the moment system (15) with closure relation (18) corresponds to a symmetric hyperbolic system, we first reformulate (15) in terms of the Lagrange multipliers. The constraints in (18) imply that  $\boldsymbol{\mu} = \langle \mathbf{m} \mathcal{B} \psi(\boldsymbol{\lambda} \cdot \mathbf{m}) \rangle$ . Hence, we have

$$\partial_t \boldsymbol{\mu} = \partial_t \langle \mathbf{m} \mathcal{B} \psi(\boldsymbol{\lambda} \cdot \mathbf{m}) \rangle = \langle \mathbf{m} \psi(\boldsymbol{\lambda} \cdot \mathbf{m}) \partial_t \mathcal{B} \rangle + \mathbf{A}_0(\boldsymbol{\lambda}) \partial_t \boldsymbol{\lambda} \quad (24)$$

with  $\mathbf{A}_0(\boldsymbol{\lambda}) = \langle \mathbf{m} \otimes \mathbf{m} \mathcal{B} \psi'(\boldsymbol{\lambda} \cdot \mathbf{m}) \rangle$ . For the flux terms, we obtain

$$\partial_{x_i} \langle v_i \mathbf{m} \mathcal{B} \psi(\boldsymbol{\lambda} \cdot \mathbf{m}) \rangle = \langle \mathbf{m} \psi(\boldsymbol{\lambda} \cdot \mathbf{m}) v_i \partial_{x_i} \mathcal{B} \rangle + \mathbf{A}_i(\boldsymbol{\lambda}) \partial_{x_i} \boldsymbol{\lambda} \quad (25)$$

with  $\mathbf{A}_i(\boldsymbol{\lambda}) = \langle v_i \mathbf{m} \otimes \mathbf{m} \mathcal{B} \psi'(\boldsymbol{\lambda} \cdot \mathbf{m}) \rangle$ . By virtue of (24) and (25), the moment system can be recast as the following quasi-linear system for the Lagrange multipliers:

$$\mathbf{A}_0(\boldsymbol{\lambda}) \frac{\partial \boldsymbol{\lambda}}{\partial t} + \sum_{i=1}^D \mathbf{A}_i(\boldsymbol{\lambda}) \frac{\partial \boldsymbol{\lambda}}{\partial x_i} = \mathbf{s}(\boldsymbol{\lambda}) \quad (26)$$

with

$$\mathbf{s}(\boldsymbol{\lambda}) = \langle \mathbf{m} \mathcal{C}(\mathcal{B} \psi(\boldsymbol{\lambda} \cdot \mathbf{m})) \rangle - \langle \mathbf{m} \psi(\boldsymbol{\lambda} \cdot \mathbf{m}) (\partial_t \mathcal{B} + v_i \partial_{x_i} \mathcal{B}) \rangle \quad (27)$$

System (26) is symmetric hyperbolic if  $\mathbf{A}_0, \mathbf{A}_1, \dots, \mathbf{A}_D$  are symmetric and  $\mathbf{A}_0$  is positive definite. The symmetry of  $\mathbf{A}_0, \mathbf{A}_1, \dots, \mathbf{A}_D$  is evident. To corroborate the positive definiteness of  $\mathbf{A}_0$ , we note that for any  $M$ -vector  $\boldsymbol{\theta}$  there holds

$$\boldsymbol{\theta} \cdot \mathbf{A}_0(\boldsymbol{\lambda}) \boldsymbol{\theta} = \langle (\boldsymbol{\theta} \cdot \mathbf{m})^2 \mathcal{B} \psi'(\boldsymbol{\lambda} \cdot \mathbf{m}) \rangle \geq 0 \quad (28)$$

The inequality holds because each of the factors  $(\boldsymbol{\theta} \cdot \mathbf{m})^2$ ,  $\mathcal{B}$  and  $\psi'(\boldsymbol{\lambda} \cdot \mathbf{m})$  is non-negative. For  $\boldsymbol{\theta} \neq 0$ , the inequality in (28) is strict, because the roots of the polynomial  $\boldsymbol{\theta} \cdot \mathbf{m}(\mathbf{v})$  are confined to a set of measure zero,  $\mathcal{B}$  is strictly positive by assumption, and  $\psi'$  is strictly positive on  $(\varphi'(0), \infty)$ . The matrix  $\mathbf{A}_0(\boldsymbol{\lambda})$  is therefore indeed positive definite. By virtue of its quasi-linear symmetric hyperbolicity, the system (26) is linearly well posed. Moreover, under suitable conditions on the initial data, local-in-time existence of solutions can be established [32]. It is to be noted that the term corresponding to  $\partial_t \mathcal{B} + v_i \partial_{x_i} \mathcal{B}$  in the production term according to (27) can cause blow up of solutions to the hyperbolic system (26) in the limit  $t \rightarrow \infty$ . Hence, the hyperbolic character of (15) with closure relation (18) ensures stability of solutions only in finite time. If the background distribution  $\mathcal{B}$  is selected such that  $\partial_t \mathcal{B} + v_i \partial_{x_i} \mathcal{B}$  vanishes, then the production term exhibits the usual dissipation properties corresponding to the collision operator, and the stability provided by the symmetric-hyperbolic character of the equations extends to the ad-infinitum limit.

The moment system (15) can alternatively be construed as a Galerkin subspace approximation of the Boltzmann equation in renormalized form; see also [27]. This Galerkin-approximation interpretation can for instance prove useful in constructing error estimates for (15) and in deriving structural properties. To elucidate the Galerkin form of (15), we define the renormalization map  $\beta : \mathcal{M} \rightarrow \mathcal{F}$  according to  $\beta(g) = \mathcal{B} \psi(g)$  and we observe that for all  $\boldsymbol{\mu} \in \mathbb{R}^M$  there exists a corresponding element  $g \in \mathcal{M}$  such that  $\mathcal{F}(\boldsymbol{\mu}) = \beta(g)$ . In particular,  $g = \boldsymbol{\lambda} \cdot \mathbf{m}$  with  $\boldsymbol{\lambda}$  the Lagrange multipliers associated with the constrained minimization problem (17). The moment system (15) can then be recast into the Galerkin form:

$$\begin{aligned} \text{Find } g \in \mathcal{L}((0, T) \times \Omega; \mathcal{M}) : \\ \partial_t \langle m \beta(g) \rangle + \partial_{x_i} \langle m v_i \beta(g) \rangle = \langle m \mathcal{C}(\beta(g)) \rangle \quad \forall m \in \mathcal{M} \text{ a.e. } (t, \mathbf{x}) \in (0, T) \times \Omega \end{aligned} \quad (29)$$

where  $\mathcal{L}((0, T) \times \Omega; \mathcal{M})$  represents a suitable vector space of functions from  $(0, T) \times \Omega$  into  $\mathcal{M}$ .

#### 4. Spatial discontinuous Galerkin finite-element approximation

For the discretization of (15) with respect to the position dependence, we consider the discontinuous Galerkin finite-element method [33]. Let  $\mathcal{H} := \{h_1, h_2, \dots\} \subset \mathbb{R}_{>0}$  denote a strictly decreasing sequence of mesh parameters whose only accumulation point is 0. Consider a corresponding mesh sequence  $\mathcal{T}^{\mathcal{H}}$ , viz., a sequence of covers of the domain by non-overlapping element domains  $\kappa \subset \Omega$ . We impose on  $\mathcal{T}^{\mathcal{H}}$  the standard conditions of regularity, shape-regularity and quasi-uniformity with respect to  $\mathcal{H}$ ; see, for instance, [33] for further details. For any  $h \in \mathcal{H}$ , we indicate by  $V^{h,p}(\Omega)$  the DG finite-element approximation space of discontinuous element-wise  $D$ -variate polynomials of degree  $\leq p$ :

$$V^{h,p}(\Omega) = \{v \in L^2(\Omega) : v|_{\kappa}(\mathbf{x}) \in \text{span}\{x_1^{p_1} x_2^{p_2} \cdots x_D^{p_D}\}, p_1 + \cdots + p_D \leq p\} \quad (30)$$

The  $p$ -dependence of  $V^{h,p}$  is contextual and will generally be suppressed. We denote by

$$V^h(\Omega, \mathcal{M}) = V^h(\Omega) \times \mathcal{M} = \{\lambda_1 m_1 + \cdots + \lambda_M m_M : \lambda_i \in V^h(\Omega)\} \quad (31)$$

the extension of  $V^h$  to  $\mathcal{M}$ -valued functions.

To facilitate the presentation of the DGFEM formulation, we introduce some further notational conventions. For any  $h \in \mathcal{H}$ , we indicate by  $\mathcal{I}^h = \{\text{int}(\partial\kappa \cap \partial\hat{\kappa}) : \kappa, \hat{\kappa} \in \mathcal{T}^h, \kappa \neq \hat{\kappa}\}$  the collection of inter-element edges, by  $\mathcal{B}^h = \{\text{int}(\partial\kappa \cap \partial\Omega) : \kappa \in \mathcal{T}^h\}$  the collection of boundary edges and by  $\mathcal{S}^h = \mathcal{B}^h \cup \mathcal{I}^h$  their union. With every edge we associate a unit normal vector  $\boldsymbol{\nu}_e$ . The orientation of  $\boldsymbol{\nu}_e$  is arbitrary except on boundary edges where  $\boldsymbol{\nu}_e = \mathbf{n}|_e$ . For all interior edges, let  $\kappa_{\pm}^e \in \mathcal{T}^h$  be the two elements adjacent to the edge  $e$  such that the orientation of  $\boldsymbol{\nu}_e$  is exterior to  $\kappa_+$ . We define jump  $[[\cdot]]_e$  and mean  $\{\{\cdot\}\}_e$  operators according to:

$$[[v]]_e = \begin{cases} (v_+ - v_-) & \text{if } e \in \mathcal{I}^h \\ v_+ & \text{if } e \in \mathcal{B}^h \end{cases} \quad \{\{v\}\}_e = \begin{cases} (v_+ + v_-)/2 & \text{if } e \in \mathcal{I}^h \\ v_+ & \text{if } e \in \mathcal{B}^h \end{cases} \quad (32)$$

where  $v_+$  and  $v_-$  refer to the restriction of the traces of  $v|_{\kappa_+}$  and  $v|_{\kappa_-}$  to  $e$ . To derive the DG formulation of the closed moment system (15), we note that for any  $\mathbf{w} \in [V^h(\Omega)]^M$  there holds:

$$\sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \mathbf{w} \cdot \partial_t \boldsymbol{\mu} + \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \mathbf{w} \cdot \partial_{x_i} \langle v_i \mathbf{m} \mathcal{F}(\boldsymbol{\mu}) \rangle = \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \mathbf{w} \cdot \langle \mathbf{m} \mathcal{C}(\mathcal{F}(\boldsymbol{\mu})) \rangle. \quad (33)$$

The second term in the left member of (33) can be recast into

$$\begin{aligned} \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \mathbf{w} \cdot \partial_{x_i} \langle v_i \mathbf{m} \mathcal{F} \rangle &= \sum_{\kappa \in \mathcal{T}^h} \int_{\partial\kappa} \mathbf{w} \cdot \langle v_i \mathbf{m} \mathcal{F} \rangle \nu_i^{\kappa} - \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \partial_{x_i} \mathbf{w} \cdot \langle v_i \mathbf{m} \mathcal{F} \rangle \\ &= \sum_{e \in \mathcal{S}^h} \int_e [[\mathbf{w}]] \cdot \langle v_{\nu} \mathbf{m} \{\{\mathcal{F}\}\} \rangle + \sum_{e \in \mathcal{I}^h} \int_e \{\{\mathbf{w}\}\} \cdot [[\langle v_{\nu} \mathbf{m} \mathcal{F} \rangle]] - \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \partial_{x_i} \mathbf{w} \cdot \langle v_i \mathbf{m} \mathcal{F} \rangle \end{aligned} \quad (34)$$

with  $\boldsymbol{\nu}^{\kappa}$  the exterior unit normal vector to  $\partial\kappa$  and  $v_{\nu} = \mathbf{v} \cdot \boldsymbol{\nu}$ . The first identity in (34) follows from the product rule and integration by parts. The second identity results from a rearrangement of terms. Implicit in the identities in (34) is the assumption that  $\mathcal{F}$  is sufficiently smooth within the elements to permit integration by parts and define traces on  $\partial\kappa$ . If, moreover,  $\langle v_{\nu} \mathbf{m} \mathcal{F} \rangle$  is continuous across the interior edges, then the second term in the right member of (34) can be removed, and  $\{\{\mathcal{F}(\boldsymbol{\mu})\}\}$  in the first term can be replaced by any  $\hat{\mathcal{F}}(\boldsymbol{\mu}_+, \boldsymbol{\mu}_-)$  subject to the consistency condition  $\hat{\mathcal{F}}(\boldsymbol{\mu}, \boldsymbol{\mu}) = \mathcal{F}(\boldsymbol{\mu})$ . On boundary edges, the external moment vector  $\boldsymbol{\mu}_-$  is to be conceived of as exogenous data in accordance with boundary conditions. It then holds that

$$[\partial_t \boldsymbol{\mu}, \mathbf{w}] + a(\boldsymbol{\mu}; \mathbf{w}) = s(\boldsymbol{\mu}; \mathbf{w}) \quad (35)$$

with

$$[\partial_t \boldsymbol{\mu}, \mathbf{w}] = \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \mathbf{w} \cdot \partial_t \boldsymbol{\mu} \quad (36)$$

$$a(\boldsymbol{\mu}; \mathbf{w}) = \sum_{e \in \mathcal{S}^h} \int_e [[\mathbf{w}]] \cdot \langle v_{\nu} \mathbf{m} \hat{\mathcal{F}}(\boldsymbol{\mu}_+, \boldsymbol{\mu}_-) \rangle - \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \partial_{x_i} \mathbf{w} \cdot \langle v_i \mathbf{m} \mathcal{F} \rangle \quad (37)$$

$$s(\boldsymbol{\mu}; \mathbf{w}) = \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \mathbf{w} \cdot \langle \mathbf{m} \mathcal{C}(\mathcal{F}(\boldsymbol{\mu})) \rangle \quad (38)$$

is consistent with (15) in the sense that any solution to (15) that is sufficiently regular in the aforementioned sense satisfies (35) for all  $\mathbf{w} \in [V^h(\Omega)]^M$ . The DGFEM approximation of (15) is obtained by replacing  $\boldsymbol{\mu}$  in (35) by an approximation  $\boldsymbol{\mu}^h$  in  $\mathcal{L}(0, T; [V^h(\Omega)]^M)$  according to:

$$\text{Find } \boldsymbol{\mu}^h \in \mathcal{L}(0, T; [V^h(\Omega)]^M) : \quad [\partial_t \boldsymbol{\mu}^h, \mathbf{w}] + a(\boldsymbol{\mu}^h; \mathbf{w}) = s(\boldsymbol{\mu}^h; \mathbf{w}) \quad \forall \mathbf{w} \in [V^h(\Omega)]^M \quad (39)$$

a.e.  $t \in (0, T)$

The edge distributions  $\hat{\mathcal{F}}$  in (37) must be constructed such that the consistency condition  $\hat{\mathcal{F}}(\boldsymbol{\mu}, \boldsymbol{\mu}) = \mathcal{F}(\boldsymbol{\mu})$  holds for all  $\boldsymbol{\mu} \in \mathbb{R}^M$  and that the formulation is stable in some appropriate sense. We propose the upwind edge distribution:

$$\hat{\mathcal{F}}(\boldsymbol{\mu}_+, \boldsymbol{\mu}_-) = \begin{cases} \mathcal{F}(\boldsymbol{\mu}_+) & \text{if } v_\nu > 0 \\ \mathcal{F}(\boldsymbol{\mu}_-) & \text{if } v_\nu \leq 0 \end{cases} \quad (40)$$

We will show that this choice of the edge distribution ensures that the DGFEM formulation (39) is entropy stable in the sense that the entropy-dissipation property (21) of the moment system (15) with closure relation (18) transfers to the DGFEM formulation. To facilitate the derivation of the entropy-dissipation property, we first recast (39) into the equivalent form of a DGFEM/moment Galerkin approximation of the renormalized Boltzmann equation according to (29):

$$\begin{aligned} \text{Find } g^h \in \mathcal{L}(0, T; V^h(\Omega, \mathcal{M})) : \\ \int_{\Omega} \langle w \partial_t \beta(g^h) \rangle + \sum_{e \in \mathcal{S}^h} \int_e \langle v_\nu \llbracket w \rrbracket \hat{\beta}(g_+^h, g_-^h) \rangle - \int_{\Omega} \langle (\partial_{x_i} w) v_i \beta(g^h) \rangle = \int_{\Omega} \langle w \mathcal{C}(\beta(g^h)) \rangle \\ \forall w \in V^h(\Omega, \mathcal{M}) \text{ a.e. } t \in (0, T) \end{aligned} \quad (41)$$

Our objective is to show that if  $g^h$  satisfies (41), then (21) holds with  $\mathcal{F}$  replaced by  $\beta(g^h)$ . Note that in the boundary integral in (21),  $\mathcal{F}$  must be replaced by  $\hat{\beta}(g_+^h, g_-^h)$ . To condense the presentation, we assume a-priori that the background distribution  $\mathcal{B}$  is uniform in space and time. Similar arguments as in Section 3 then convey

$$\frac{d}{dt} \int_{\Omega} \langle \mathcal{B} \varphi(\beta^h / \mathcal{B}) \rangle = \int_{\Omega} \langle g^h \partial_t \beta^h \rangle + \int_{\Omega} \int_{\text{supp}^c(\beta^h)} (\varphi'(0) - g^h) \partial_t \beta^h \quad (42)$$

with the abridged notation  $\beta^h := \beta(g^h)$ . Noting that  $g^h(t)$  resides in  $V^h(\Omega, \mathcal{M})$  a.e.  $t \in (0, T)$ , we can apply Galerkin orthogonality according to (41) to recast the first term in (42) into:

$$\int_{\Omega} \langle g^h \partial_t \beta^h \rangle = - \sum_{e \in \mathcal{S}^h} \int_e \langle v_\nu \llbracket g^h \rrbracket \hat{\beta}^h \rangle + \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \langle (\partial_{x_i} g^h) v_i \beta^h \rangle + \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \langle g^h \mathcal{C}(\beta^h) \rangle \quad (43)$$

with  $\hat{\beta}^h := \hat{\beta}(g_+^h, g_-^h)$ . The second term in (43) can be recast into

$$\sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \langle (\partial_{x_i} g^h) v_i \beta^h \rangle = \sum_{e \in \mathcal{S}^h} \langle v_\nu \llbracket g^h \beta^h \rrbracket \rangle - \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \langle g^h v_i \partial_{x_i} \beta^h \rangle \quad (44)$$

The second term in the right member of (44) can in turn be reformulated as

$$\begin{aligned} \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \langle g^h v_i \partial_{x_i} \beta^h \rangle &= \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \langle \varphi'(\beta^h / \mathcal{B}) v_i \partial_{x_i} \beta^h \rangle - \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \int_{\text{supp}^c(\beta^h)} (\varphi'(0) - g^h) v_i \partial_{x_i} \beta^h \\ &= \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \partial_{x_i} \langle v_i \mathcal{B} \varphi(\beta^h / \mathcal{B}) \rangle - \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \int_{\text{supp}^c(\beta^h)} (\varphi'(0) - g^h) v_i \partial_{x_i} \beta^h \\ &= \sum_{e \in \mathcal{S}^h} \langle v_\nu \llbracket \mathcal{B} \varphi(\beta^h / \mathcal{B}) \rrbracket \rangle - \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \int_{\text{supp}^c(\beta^h)} (\varphi'(0) - g^h) v_i \partial_{x_i} \beta^h \end{aligned} \quad (45)$$

Collecting the results in Equations (42)-(45) and recalling that  $g = \varphi'(\beta(g)/\mathcal{B})$  on the support of  $\beta$ , we obtain

$$\frac{d}{dt} \int_{\Omega} \langle \mathcal{B} \varphi(\beta^h / \mathcal{B}) \rangle + \int_{\partial\Omega} \langle v_n \mathcal{B} \varphi(\hat{\beta}^h / \mathcal{B}) \rangle = T_{\mathcal{S}^h} + T_{\mathcal{T}^h} \quad (46)$$



with

$$T_{S^h} = \sum_{e \in S^h} \left\langle v_\nu \llbracket \varphi'(\beta^h/\mathcal{B})(\beta^h - \hat{\beta}^h) - (\mathcal{B}\varphi(\beta^h/\mathcal{B}) - \mathcal{B}\varphi(\hat{\beta}^h/\mathcal{B})) \rrbracket \right\rangle \\ - \sum_{e \in S^h} \left\llbracket \int_{\text{supp}^c(\beta^h)} v_\nu(\varphi'(0) - g^h)(\beta^h - \hat{\beta}^h) \right\rrbracket \quad (47)$$

$$T_{\mathcal{T}^h} = \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \langle \varphi'(\beta^h/\mathcal{B}) \mathcal{C}(\beta^h) \rangle + \sum_{\kappa \in \mathcal{T}^h} \int_{\kappa} \int_{\text{supp}^c(\beta^h)} (\varphi'(0) - g^h)(\partial_t \beta^h + v_i \partial_{x_i} \beta^h - \mathcal{C}(\beta^h)) \quad (48)$$

Let us note that  $\llbracket \mathcal{B}\varphi(\hat{\beta}^h/\mathcal{B}) \rrbracket$  in (47) vanishes on interior edges by virtue of the continuity of  $\hat{\beta}^h$ , and that its aggregated contribution coincides with the boundary integral in the left member of (46). Similar arguments as in Section 3 convey that  $T_{\mathcal{T}^h} \leq 0$ ; cf. (23). To assess the contribution of  $T_{S^h}$ , we recall that  $\hat{\beta}^h = \beta_+^h$  (resp.  $\hat{\beta}^h = \beta_-^h$ ) if  $v_\nu > 0$  (resp.  $v_\nu \leq 0$ ). For  $v_\nu > 0$  it therefore follows that

$$\llbracket \varphi'(\beta^h/\mathcal{B})(\beta^h - \hat{\beta}^h) - (\mathcal{B}\varphi(\beta^h/\mathcal{B}) - \mathcal{B}\varphi(\hat{\beta}^h/\mathcal{B})) \rrbracket = -(\varphi'(\beta_-^h/\mathcal{B})(\beta_-^h - \hat{\beta}^h) - (\mathcal{B}\varphi(\beta_-^h/\mathcal{B}) - \mathcal{B}\varphi(\hat{\beta}^h/\mathcal{B}))) \quad (49)$$

From the convexity of  $\mathcal{B}\varphi(\cdot/\mathcal{B})$  we infer that for any  $\beta, \hat{\beta} \in \mathbb{R}$ :

$$\varphi'(\beta/\mathcal{B})(\beta - \hat{\beta}) \geq \mathcal{B}\varphi(\beta/\mathcal{B}) - \mathcal{B}\varphi(\hat{\beta}/\mathcal{B}) \quad (50)$$

Hence, for  $v_\nu > 0$  the jump term in the left member of (49) is non-positive. Similarly, one can infer that the jump term is non-negative if  $v_\nu \leq 0$ . The first term in  $T_{S^h}$  is therefore non-positive. The contribution of each edge to the second term in  $T_{S^h}$  can be decomposed as:

$$\left\llbracket \int_{\text{supp}^c(\beta^h)} v_\nu(\varphi'(0) - g^h)(\beta^h - \hat{\beta}^h) \right\rrbracket = \int_{\text{supp}^c(\beta_+^h), v_\nu \leq 0} v_\nu(\varphi'(0) - g_+^h)(\beta_+^h - \beta_-^h) \\ - \int_{\text{supp}^c(\beta_-^h), v_\nu > 0} v_\nu(\varphi'(0) - g_-^h)(\beta_-^h - \beta_+^h) \quad (51)$$

It generally holds that  $\varphi'(0) \geq g_\pm^h$ ; cf. (19). In the first term in (51) the domain of integration is a subset of  $\text{supp}^c(\beta_+^h)$  and accordingly  $\beta_+^h$  vanishes. It then follows from  $v_\nu \leq 0$  and  $\beta_-^h \geq 0$  that this term is non-negative. In a similar manner, it can be shown that the contribution of the second term in (51) is also non-negative. Both terms in (47) are therefore non-positive.

## 5. Numerical results

To illustrate the properties of the proposed discontinuous Galerkin finite-element moment method (41), or equivalently (39), we present numerical experiments for a one-dimensional shock-tube problem. Before doing so, the problem specification and moment-system approximation must be completed by specifying the collision operator and closure relation. We restrict ourselves here to the standard BGK collision operator [31], viz.  $\mathcal{C}(f) = -\tau^{-1}(f - \mathcal{M}_f)$  with  $\mathcal{M}_f$  the local Maxwellian with the same invariant moments as  $f$  and  $\tau^{-1}$  a relaxation rate. We adopt the relaxation parameter in accordance with the hard-sphere collision process of Bird [17]:  $\tau = (5\lambda/16)(2\pi\rho/p)^{1/2}$  with  $\lambda$  the mean free path. We identify the Knudsen number with the mean free path divested of its units. The closure relation is defined by the renormalization map:

$$\beta(g) = \mathcal{B} \left( 1 + \frac{g}{N} \right)_+^N, \quad (52)$$

where  $(\cdot)_+ = \frac{1}{2}(\cdot) + \frac{1}{2}|\cdot|$  is the non-negative part of a function extended by 0 and  $N \in \mathbb{N}$ ; see [27] for further details on this closure relation. The corresponding  $\varphi$ -divergence based entropy is  $f \mapsto \langle \mathcal{B}\varphi(f/\mathcal{B}) \rangle$  with  $\varphi$  according to (16), i.e. (52) corresponds to the solution of the constrained entropy-minimization problem (17) with  $\varphi$  from (16). In [27] it is shown that the minimization problem in (17) is well-posed for this entropy. Moreover,  $\langle \mathcal{B}\varphi(f/\mathcal{B}) \rangle$  corresponds to an entropy for the BGK operator in compliance with the dissipation relation (8); see also [27]. In the sequel, we set  $N = 2$ .

We regard Sod's shock-tube problem on a spatial domain  $x \in (-1, 1)$  and a time interval  $t \in (0, 0.1)$ . This test case concerns a Riemann problem with discontinuous initial data corresponding to a piecewise uniform Maxwellian distribution:

$$f(0, x, v) = \begin{cases} f_L(v) = \mathcal{M}_{(1,0,1)}(v) & \text{for } x \leq 0 \\ f_R(v) = \mathcal{M}_{(\frac{1}{2},0,\frac{2}{5})}(v) & \text{for } x > 0 \end{cases} \quad (53)$$

cf. (12); see [34]. Two different approximations are considered. For the first approximation we regard a uniform background distribution  $\mathcal{B} = \mathcal{M}_{(1,0,1)}$ . In the second approximation the background distribution corresponds to the solution of the Vlasov equation with initial data (53), locally regularized near  $x = 0$ :

$$\partial_t \mathcal{B} + \partial_x v \mathcal{B} = 0 \quad (54a)$$

$$\mathcal{B}(0, x, v) = \mathcal{B}_0(x, v) = \begin{cases} \mathcal{M}_{(1,0,1)}(v) & \text{for } x \leq x_0 \\ \frac{x_1 - x}{x_1 - x_0} \mathcal{M}_{(1,0,1)}(v) + \frac{x - x_0}{x_1 - x_0} \mathcal{M}_{(\frac{1}{2},0,\frac{2}{5})}(v) & \text{for } x_0 < x < x_1 \\ \mathcal{M}_{(\frac{1}{2},0,\frac{2}{5})}(v) & \text{for } x \geq x_1 \end{cases} \quad (54b)$$

with  $x_1 = -x_0 = \frac{1}{64}$ . The regularization near  $x = 0$  serves to avoid complications related to discontinuities in the background distribution; see Section 3. The solution to (54) is given by  $\mathcal{B}(t, x, v) = \mathcal{B}_0(x - vt, v)$ . The parameters  $\rho, p$  in the relaxation parameter  $\tau$  are determined from  $f_L$  in the initial distribution (53).

We restrict ourselves here to finite-element approximation spaces of polynomial degree  $p = 0$ , i.e. element-wise constant approximations in position dependence. For the time-integration procedure, we apply the forward Euler scheme with a time step that is sufficiently small to render the numerical results essentially independent of the time step.

Figure 1 displays the evolution of the entropy  $\int_{\Omega} \langle \mathcal{B} \varphi(\mathcal{F}/\mathcal{B}) \rangle$  for the 5-moment system, i.e. for  $\mathcal{M}$  spanned by  $\{1, v, \dots, v^4\}$ , for both the uniform and non-uniform background distributions, for Knudsen numbers  $\text{Kn} \in 2.3 \times 10^{\{-2,0,+1\}}$ , and for uniform meshes with mesh width  $h \in 2^{\{-7,\dots,-10\}}$ . The results in Figure 1 corroborate the entropy-dissipation property of the DGFEM moment method. Observing that the results for Knudsen numbers  $\text{Kn} \in 2.3 \times 10^{\{0,+1\}}$  are nearly identical, we infer that in the corresponding highly rarefied and transition regimes the entropy dissipation is dominated by the dissipation induced by the discontinuities in the DG approximation, for the considered sequence of finite-element spaces. This is also evident from the deviation between the graphs for distinct  $h$ . In the near continuum regime  $\text{Kn} = 2.3 \times 10^{-2}$  and on sufficiently fine meshes, the effect of the discontinuities on the entropy dissipation is less pronounced.

Figure 2 presents the normalized heat flux according to

$$q^* = \rho^{1/2} p^{-3/2} q \quad (55)$$

with  $\rho = \langle f \rangle$  as density,  $p = \frac{1}{2} \langle (u - v)^2 f \rangle$  as pressure,  $q = \langle (v - u)^3 f \rangle$  as heat flux and  $u = \langle v f \rangle / \langle f \rangle$  as the macroscopic velocity, in the highly-rarefied regime ( $\text{Kn} = 2.3 \times 10^{+1}$ ) at  $t = 1/10$ , extracted from the DGFEM moment-method approximation for  $M \in \{5, 7, 9\}$  and mesh width  $h = 2^{-10}$ , for both the uniform and non-uniform background distribution. Let us note that heat transfer is a typical non-equilibrium effect that is not represented by the conventional continuum model of gas dynamics, viz. the Euler equations with ideal-gas closure. The results in Figure 2 convey that the approximation of the heat flux for the uniform background distribution is still relatively sensitive to the number of moments that is applied in the approximation. By virtue of the high Knudsen number, the background distribution (54) derived from the Vlasov solution yields an accurate approximation to the actual solution of the Boltzmann equation, which translates into very high accuracy of the corresponding heat-flux approximation, virtually independent of the number of moments. It is to be noted, however, that the accuracy of the heat-flux approximation pertaining to the uniform background distribution in Figure 2 (*left*) also compares favorably to corresponding results in the literature for other closure relations; cf. for instance [34].

## 6. Conclusion

We have presented a Galerkin approximation method for the Boltzmann equation based on the combination of moment-system approximation in velocity dependence and discontinuous-Galerkin finite-element approximation in position dependence. For the moment systems, we considered a closure relation that derives from the minimization of a relative entropy corresponding to a  $\varphi$ -divergence. The background measure in the relative

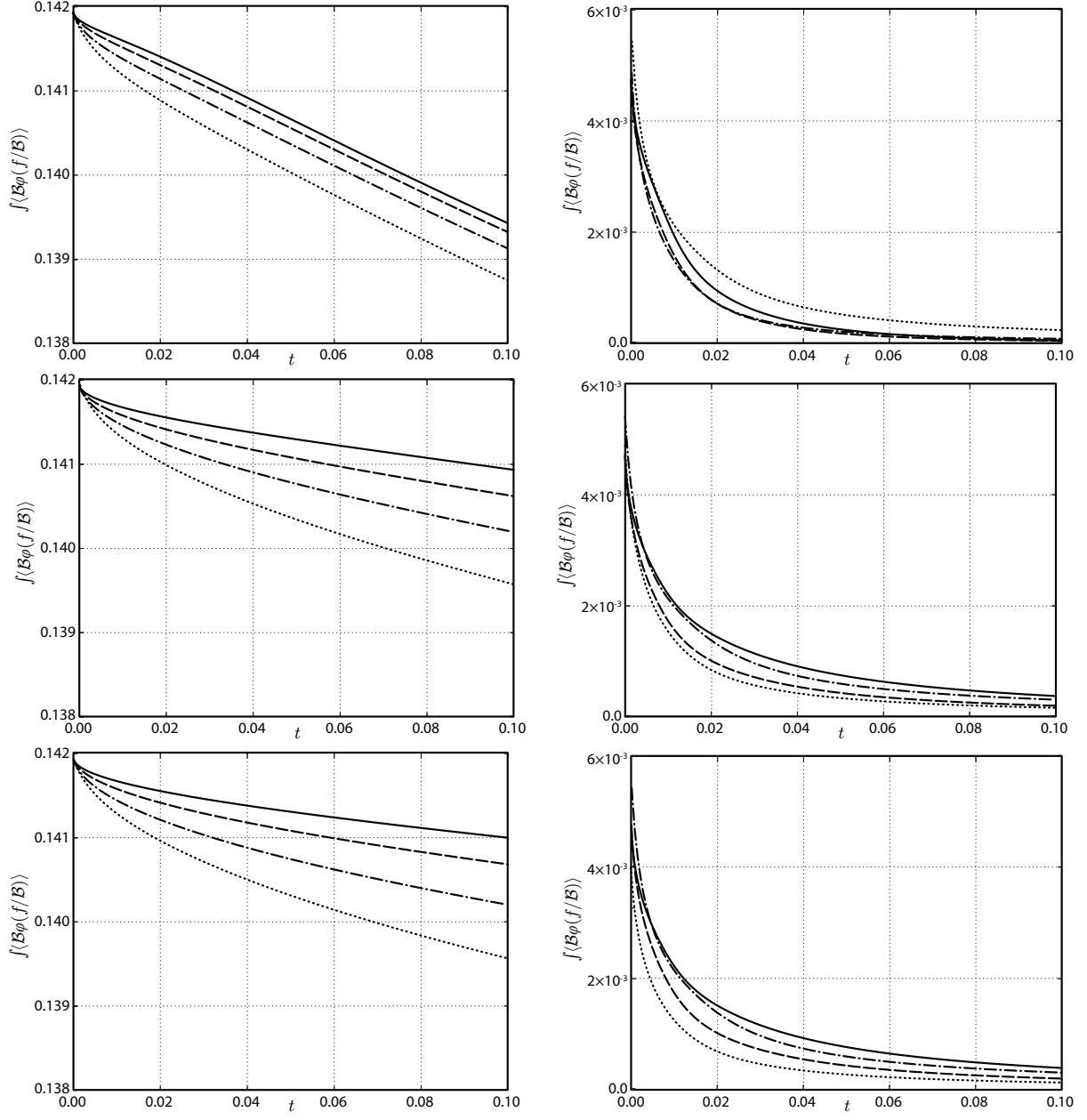


Figure 1: Evolution of the entropy  $\int_{\Omega} \langle \mathcal{B} \varphi(\mathcal{F}/\mathcal{B}) \rangle$  for the DGFEM 5-moment approximation, for the uniform background distribution  $\mathcal{B} = \mathcal{M}_{(1,0,1)}$  (left) and the non-uniform background distribution in (54) (right), for Knudsen numbers  $\text{Kn} = 2.3 \times 10^{-2}$  (top),  $\text{Kn} = 2.3 \times 10^0$  (center) and  $\text{Kn} = 2.3 \times 10^{+1}$  (bottom), and for uniform meshes with mesh width  $h = 2^{-7}$  ( $\cdots$ ),  $h = 2^{-8}$  ( $-\cdot-$ ),  $h = 2^{-9}$  ( $- -$ ) and  $h = 2^{-10}$  ( $-$ ).

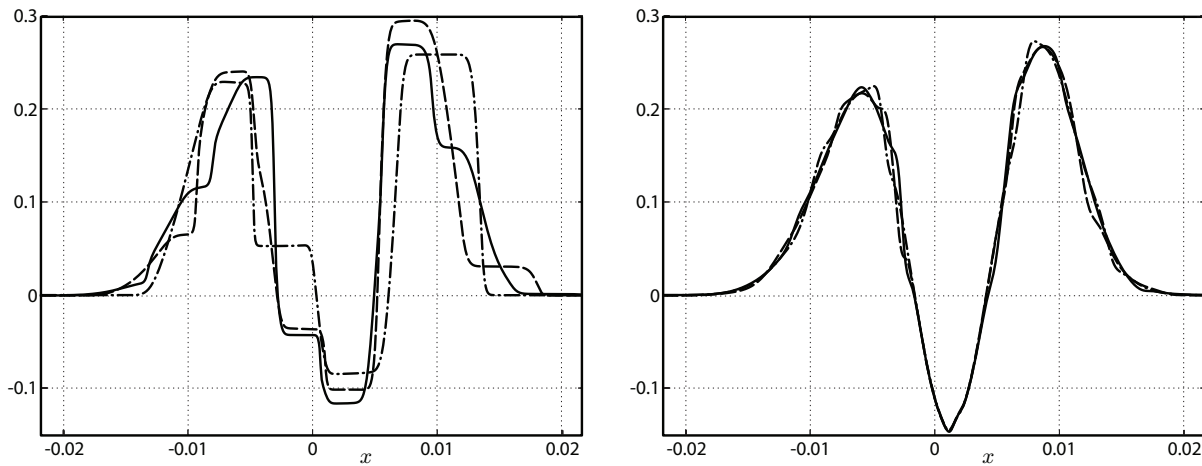


Figure 2: Heat flux  $q^*$  according to (55) for  $\text{Kn} = 2.3 \times 10^{+1}$  at  $t = 1/10$ , for the uniform background distribution  $\mathcal{B} = \mathcal{M}_{(1,0,1)}$  (left) and the non-uniform background distribution in (54) (right), for  $M = 5$  ( $\cdots$ ),  $M = 7$  ( $-\cdots-$ ) and  $M = 9$  ( $\text{---}$ ) moments and mesh width  $h = 2^{-10}$ .

entropy then appears as a factor in the approximation of the one-particle marginal. We established that for suitable background measures, the moment systems retain the fundamental structural properties of the underlying Boltzmann equation, viz., conservation of mass, momentum and energy, Galilean invariance and dissipation of a relative entropy. Moreover, the divergence-based closure leads to a hierarchy of tractable symmetric hyperbolic systems. The moment systems can alternatively be conceived of as Galerkin approximations of the Boltzmann equation in renormalized form.

For the discretization of the moment systems in position dependence, we considered the discontinuous Galerkin finite-element method. The combined DGFEM moment method can be construed as a Galerkin finite-element approximation of the Boltzmann equation in renormalized form, based on a tensor-product approximation space composed of the DGFEM approximation space in position dependence and global polynomials in velocity dependence. We introduced a new class of numerical flux functions for the combined DGFEM moment method. This new numerical flux function appears naturally in the setting of the renormalized Boltzmann equation as the upwind distribution at the inter-element interfaces. We established that this upwind flux renders the DGFEM moment method entropy stable, i.e. the entropy-dissipation property of the moment systems transfers to the DGFEM formulation.

Numerical results were presented for a one-dimensional shock-tube problem in the highly-rarefied, intermediate and near-continuum regimes. We considered two different approximations: one based on a background distribution corresponding to a uniform Maxwellian, and one based on a background distribution corresponding to a solution to the Vlasov equation. The numerical results confirm the entropy-dissipation property of the DGFEM moment method for both types of approximation. In the highly rarefied regime, the approximation based on the Vlasov background distribution provides excellent approximations of the heat flux, essentially independent of the number of moments. The approximation corresponding to the uniform background distribution yields very accurate results for the heat flux, that compare favorably to corresponding results in the literature for other closure relations.

## References

- [1] C. Bardos, F. Golse, D. Levermore, Fluid dynamic limits of kinetic equations. I. Formal derivations, J. Stat. Phys. 63 (1) (1991) 323–344.
- [2] R. Esposito, J. Lebowitz, R. Marra, Hydrodynamic limit of the stationary Boltzmann equation in a slab, Communications in Mathematical Physics 160 (1) (1994) 49–80.
- [3] F. Golse, L. Saint-Raymond, The Navier–Stokes limit of the Boltzmann equation for bounded collision kernels, Inventiones Mathematicae 155 (1) (2004) 81–161.

- [4] P.-L. Lions, N. Masmoudi, From the Boltzmann Equations to the Equations of Incompressible Fluid Mechanics, I, Arch. Rational Mech. Anal. 158 (3) (2001) 173–193.
- [5] P.-L. Lions, N. Masmoudi, From the Boltzmann Equations to the Equations of Incompressible Fluid Mechanics, II, Arch. Rational Mech. Anal. 158 (3) (2001) 195–211.
- [6] C. Levermore, N. Masmoudi, From the Boltzmann Equation to an Incompressible Navier–Stokes–Fourier System, Arch. Rational Mech. Anal. 196 (2010) 753–809.
- [7] L. Saint-Raymond, Hydrodynamic Limits of the Boltzmann Equation, no. v. 1971 in Lecture Notes in Mathematics, Springer, 2009.
- [8] C. Cercignani, Rarefied Gas Dynamics: From Basic Concepts to Actual Calculations, Cambridge Texts in Applied Mathematics, Cambridge University Press, 2000.
- [9] H. Struchtrup, Macroscopic Transport Equations for Rarefied Gas Flows: Approximation Methods in Kinetic Theory, Interaction of Mechanics and Mathematics Series, Springer, 2005.
- [10] C. Shen, Rarefied Gas Dynamics: Fundamentals, Simulations and Micro Flows, Heat and mass transfer, Springer Berlin Heidelberg, 2005.
- [11] A. Jüngel, Transport Equations for Semiconductors, Lecture Notes in Physics, Springer, 2009.
- [12] K. Miyamoto, Plasma Physics and Controlled Nuclear Fusion, Springer, 2004.
- [13] M. Reeks, On a kinetic equation for the transport of particles in turbulent flows, Physics of Fluids A: Fluid Dynamics 3 (3) (1991) 446–456. doi:10.1063/1.858101.
- [14] M. Reeks, On the continuum equations for dispersed particles in nonuniform flows, Physics of Fluids A: Fluid Dynamics 4 (6) (1992) 1290–1303. doi:10.1063/1.858247.
- [15] M. Reeks, On the constitutive relations for dispersed particles in nonuniform flows. I: Dispersion in a simple shear flow, Physics of Fluids A: Fluid Dynamics 5 (3) (1993) 750–761. doi:10.1063/1.858658.
- [16] G. Bird, Direct Simulation and the Boltzmann Equation, Phys. Fluids 13 (1970) 2676–2681.
- [17] G. Bird, Molecular Gas Dynamics and the Direct Simulation of Gas Flows, Oxford University Press, 1994.
- [18] W. Wagner, A Convergence Proof for Bird’s Direct Simulation Monte Carlo Method for the Boltzmann Equation, Journal of Statistical Physics 66 (1992) 1011–1044.
- [19] A. Klenke, Probability Theory: A Comprehensive Course, Universitext (1979), Springer, 2008.
- [20] H. Grad, On the kinetic theory of rarefied gases, Communications on Pure and Applied Mathematics 2 (4) (1949) 331–407.
- [21] C. Levermore, Moment closure hierarchies for kinetic theories, Journal of Statistical Physics 83 (1996) 1021–1065.
- [22] L. Mátyás, Generalized Method of Moments Estimation, Themes In Modern Econometrics, Cambridge University Press, 1999.
- [23] W. Dreyer, Maximisation of the entropy in non-equilibrium, Journal of Physics A: Mathematical and General 20 (18) (1987) 6505.
- [24] I. Müller, T. Ruggeri, Extended thermodynamics, Springer tracts in natural philosophy, Springer-Verlag, 1993.
- [25] S. Kullback, R. Leibler, On Information and Sufficiency, The Annals of Mathematical Statistics 22 (1) (1951) pp. 79–86.
- [26] I. Csizsár, A class of measures of informativity of observation channels, Periodica Mathematica Hungarica 2 (1972) 191–213.

- [27] M. Abdel-Malik, E. van Brummelen, Moment closure approximations of the boltzmann equation based on  $\varphi$ -divergences, ArXiv arXiv:1503.05183 [math-ph] (2015) 1–20.
- [28] R. Courant, K. Friedrichs, Supersonic Flow and Shock Waves, Applied Mathematical Sciences, Springer New York, 1999.
- [29] E. Toro, Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction, Springer Berlin Heidelberg, 2013.
- [30] S. Ali, S. Silvey, A General Class of Coefficients of Divergence of One Distribution from Another, Journal of the Royal Statistical Society. Series B (Methodological) 28 (1) (1966) pp. 131–142.
- [31] P. Bhatnagar, E. Gross, M. Krook, A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems., Phys. Rev. 94 (511—525).
- [32] A. Majda, Compressible Fluid Flow and Systems of Conservation Laws in Several Space Dimensions, Springer, Berlin, 1984.
- [33] D. Di Pietro, A. Ern, Mathematical Aspects of Discontinuous Galerkin Methods, Springer, 2012.
- [34] J. McDonald, C. T. Groth, Towards realizable hyperbolic moment closures for viscous heat-conducting gas flows based on a maximum-entropy distribution, Continuum Mech. Thermodyn. 25 (5) (2013) 573–603. doi:10.1007/s00161-012-0252-y.